AD_____

Award Number: DAMD17-01-1-0529

TITLE: Antibody Probes to Transcript-Specific Peptides Provide a Novel Tool to Investigate the Role of Alternate Estrogen Receptor Promoter Use in Breast Cancer

PRINCIPAL INVESTIGATOR: Brian T. Pentecost, Ph.D.

CONTRACTING ORGANIZATION: Health Research, Incorporated
Rensselaer, New York 12144

REPORT DATE: July 2003

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	July 2003	Final (1 Jul 2001 - 30 Jun 2003)

4. TITLE AND SUBTITLE
Antibody Probes to Transcript-Specific Peptides Provide a Novel Tool to Investigate the Role of Alternate Estrogen Receptor Promoter Use in Breast Cancer

5. FUNDING NUMBERS
DAMD17-01-1-0529

6. AUTHOR(S)
Brian T. Pentecost, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Health Research, Incorporated
Rensselaer, New York 12144

E-Mail: Brian.Pentecost@wadsworth.org

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
Original contains color plates: All DTIC reproductions will be in black and white.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 Words)

Estrogen receptor alpha (ER) plays an important role in the development and progression of breast cancer, and it is routinely used as a marker for hormone sensitivity in breast cancer patients. Positive ER status is a useful indicator for a first-line therapy with antiestrogens. ER is expressed from at least two promoters (Fig.1). The resulting transcripts from these two promoters differ only in the non-coding region upstream of the major ER open reading frame (ORF); the ER proteins from these two promoters are identical. The proximal promoter transcript contains a 20 residue ORF which closes 52 nt upstream of the main ER ORF. Which affects expression from the downstream ER ORF. **Our central hypothesis is that the action of the proximal transcript uORF is exerted at the *translation* level.**

C-terminal truncated (Phe20 or Gly 19 to stop) prox-uORFs are highly effective and enhanced inhibitors of expression from the downstream ORF. We have data from the funded studies that supports the translational control hypothesis. The study has demonstrated consequences of alternate promoter use beyond relative strength of the alternate promoters.

14. SUBJECT TERMS
Gene expression regulation and control, estrogen receptor

15. NUMBER OF PAGES
25

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

Table of Contents

INTRODUCTION

Estrogen receptor alpha (ER) plays an important role in the development and progression of breast cancer, and it is routinely used as a marker for hormone sensitivity in breast cancer patients (1). Positive ER status is a useful indicator for a first-line therapy with antiestrogens (mainly tamoxifen) (2).

ER is expressed from at least two promoters (Fig. 1). The resulting transcripts from these two promoters differ only in the non-coding region upstream of the major ER open reading frame (ORF); the ER proteins from these two promoters are identical. The distal promoter is 2 Kbp upstream of the main translational start of the ER ORF and is spliced into nt164 of the proximal form. Both proximal and distal ER promoter transcripts have short peptide ORFs in the unique regions upstream of ER itself. The proximal promoter transcript contains a 20 residue ORF which closes 50 nt upstream of the main ER ORF. The distal promoter transcript contains two ORFs. One of these, with 18 residues, shares 5 C-terminal codons with the prox-peptide of the proximal promoter transcript.

Our preliminary data indicated that the proximal transcript uORF affected expression from the downstream ER ORF, while the equivalent distal transcript ORF did not. **Our central hypothesis is that the action of the proximal transcript uORF is exerted at the *translation* level.**

A loose consensus for the region around the eukaryotic initiator AUG was defined almost twenty years ago (3). Strong start sites (Kozak consensus) tend (95%) to have purines (normally A) at -3 position relative to the AUG, and G at +4 position (Fig. 1B). The basic scanning model of eukaryotic translation machinery presumes that the first 5' AUG in a sequence should encode the translational initiator methionine (4) and any upstream ORF should be expressed to some degree with the potential to influence downstream expression. It is well known that re-initiation is required before the downstream translational start codon can be utilized. An example of that is the *S cerevisiae* GCN4 gene (5). In addition, translational regulation of specific mRNAs has been studied and is generally exerted through *cis*-acting elements in the mRNAs. One example of that is iron regulatory proteins which inhibit ferritin translation by binding to specific 5' UTR sites (6). The studied ORFs (5 principal eukaryotic examples) lack conserved (shared) features (7) and lack homology to ER ORFs.
This project was funded as a one year concept with an additional no cost extension. The project has principally supported a graduate student, Mingfei Luo, who has carried out most of the studies.

BODY

Task 1. *Carry out mutagenesis of 20 residue upstream Open Reading Frame (uORF) of the Estrogen Receptor (ER) proximal promoter transcript. Targets will address the role of hydrophobicity and charge in the action of the peptide on translation, in addition any role for the terminal phenylalanine codon will be specifically studied as there is a similar terminal residue in an ERβ uORF. Where changes affect inhibitory action, then additional changes will be made which are equivalent in effects on peptide sequence but involve alternate codon use at the nucleic acid level. This will determine if changes are due to effects on peptide sequence rather than on RNA conformation: Studies will continue to use GFP fusions as a reporter unless results of task 3 indicate that a better signal can be obtained with intact ER.*

Models for studies: We used a Green Fluorescent Protein (GFP) reporter system (see construct summary in Fig. 2) to investigate the translational capacity of the 5' untranslated region and the AUG start regions (the first 18 amino acid (AA) residues) of the ER coding regions. The main analytical tools we used were flow-cytometry and protein immuno-blotting. We transiently transfected several ER positive and negative cells. Cell lines included ER-negative Hela and breast tumor line MDA-MB231 together with ER-positive MCF7 breast cancer cells and uterine Ishikawa tumor line ..

Our studies focus on translational controls in ER transcripts, so ER-GFP transcripts of the constructs in pEGFP-N1 were from a strong, common, CMV promoter. We did not use the ER promoters. We made ER-

GFP fusion constructs in which GFP protein was generated from the ER-GFP fusion in which the natural AUG region of the major ER ORF provides the translational start. In addition, upstream region beginning seven bases prior to the proximal uORF were included in each construct, ie. construct included the uORF and the "Kozak region" around the upstream translational start. It also enabled us to mutate several of the proximal uORF codons by changes to PCR primers. The proximal uORF is present in its natural relationship to the main ER-ORF AUG in these constructs, and has the potential to influence translational of the downstream ORF. The GFP fusion model has some advantages for our studies. GFP product can be easily assayed because it fluoresces. ER could be used as "reporter" in protein immuno-blotting analysis (8), but the GFP reporter has the advantages: 1). we can assay it in cell extracts that have ER (ER positive cell lines); 2) ER-GFP fusion is quite stable in the presence of residual estrogen in the media, we do not have to worry about the estradiol stimulated ER instability and turnover.

Plasmid DNA was introduced by lipofection. For the cells used for flow-cytometry analysis, two ug of GFP construct was used with 35 mm diameter wells, with total DNA adjusted to 4 ug using pBSK. For the cells used for protein immuno-blotting analysis, one ug of GFP construct and one ug of reference construct HEM45-GFP construct were used with 35 mm diameter wells, with total DNA adjusted to 4 ug using pBSK. After incubation and change of media, cells are harvested for flow cytometry or protein immuno-blotting analysis.

The HEM45 sequence (9) was characterized with prior BCRP support. The gene has similarity to several proteins affecting cell cycle and cell differentiation however, we were never able to affect cell growth with transfected HEM45 constructs so the HEM45 should not affect cell characteristics in the current work.. The HEM45-GFP and ER-GFP fusion proteins are easily resolved on the SDS polyacrylamide gels using electrophoresis and are detected simultaneously in blots using an antisera against the common GFP component. HEM45-GFP construct encodes a ~40K Da protein larger than ER-GFP (~22 K Da) and it is quite stable though expressed at lower levels than the unmodified GFP ORF in pEGFPN1. The ER-GFP constructs should express an ER-GFP fusion protein that is slightly larger than GFP due to the utilization of the major ER translational start 18 codons upstream of that for GFP. This was confirmed in preliminary studies.

Assays for expression studies: Cells for protein immuno-blotting analysis were lysed *in-situ* using a modified gel loading solution containing 2% SDS and 62.5 mM tris pH 6.8 but lacking Beta-mercaptoethanol and bromophenol blue. Protein concentration were determined by the BCA method (Pierce, Rockford, IL) after denaturation. Typically 20 ug of protein from each sample was analyzed on 12% SDS polyacrylamide gels and blotted to PVDF membrane (Owl scientific). GFP protein was detected by polyclonal anti-GFP antibodies (Santa Cruz Biotechnology, Santa Cruz, CA) with chemiluminescence via HRP conjugated secondary antibody.

Our primary GFP assay utilized quantitative analysis by flow cytometry. Cells were harvested by trypsinization, spun down and resuspended in PBS for analysis. Analysis was restricted to single cells; debri and clumped cells were gated out based on forward and side scatter. We also gated out untransfected cells that have a basal autofluorescence. Presented data is based on the mean of fluorescence intensity (MFI) of analyzed cultures. The arithmetic mean and standard error of grouped sample MFIs were subsequently determined and are presented as a measure of fluorescence representing GFP expression from constructs in cells. Note that the HEM45-GFP constructs were not transfected to cells that were to be processed for flow cytometry

We used site-directed mutagenesis to make point mutations within the proximal uORF. The 'QuikChange' kit from Stratagene was used, this system involves the replication of plasmids in-vitro by extension of primers containing the target regions; the product is reintroduced to E coli after elimination of the parent plasmid by digestion with a restriction enzyme that will not recognize sites in the in-vitro generated DNA due to lack of methylation. Primer sets made for the study are indicated in Fig. 3. Preliminary data available when submitting our original applications for funding showed that mutating the proximal uORF

ATG to CTG (eliminating the prox-peptide translational start) resulted in a marked increase in the signal (expression of the downstream ER-GFP ORF) compared with the constructs without the mutation. We saw the same effect during this study (Fig. 4, lane 2 vs 1, Fig. 5. lane 1 vs 5).

Results for testing uORF regulatory effects: We utilized mutagenesis of specific residues to Alanine as an approach to test the role of the regions of the uORF that are unique to the prox transcript.
We made several single amino acid substitutions (listed in Fig. 3) and found no clear cut effects on expression of the downstream ER-GFP cassette.

In an alternate approach we inserted stop codons in the uORF to test structural requirements. Note that insertion of the stop codons truncates the peptide without limited effect on the RNA/DNA structure, avoiding some of the problems of eliminating stretches of DNA, an approach we tried in studies of aim 2.

A clear result of these studies was that replacing the C-terminal Phe20 codon with a stop codon greatly decreased the ER-GFP fusion protein expression in Hela cell lines as analyzed by both flow-cytometry analysis (Fig. 4, lane 3 vs 1) and quantitative western blotting (see Fig 10 discussed late in the report) and in several ER+/- cell lines in protein immuno-blotting analysis (Fig. 5, lane 2 vs 5).

In order to further investigate the relationship between the proximal ORF and the downstream expression, we generated a double mutant construct which contained the proximal ORF translational start ATG mutated to CTG and the C terminal Phe20 mutated to Stop codon. We observed a marked increase in ER-GFP expression (Fig. 4, lane 4 vs 3, Fig. 5. lane 3 vs 2). The inhibitory effect of the modified uORF is likely related to the peptide of the uORF since knocking out the translational start knocked out inhibition.

We then mutated the next to last codon (Gly 19) to stop as the terminal Phe20 was clearly NOT the source of inhibitory action. Our data showed that the Gly 19 to Stop mutant had decreased expression compared with the basic ER-GFP construct (Fig. 4, lane 7 vs 1).. In order to further investigate the inhibitory effect of the uORF peptide, we moved the proximal ORF translational start downstream (codon 15, 10) using the construct with *both* the Phe20 to stop and no uORF ATG as the starting point. Decreased expression was observed (Fig. 4. lane 5 & 6) compared to the basic 'wild type' prox-ER GFP construct (Fig. 4, lane 1). Together the data indicate roles for multiple residues in the prox-peptide uORF. In order to rule out the possibility that the further decreased ER-GFP expression observed from Phe20 to Stop mutant was not due to two termination codons, we made another mutation construct with the Phe20 deleted (looped out). Data from flow cytometry (not shown) and immuno blot (Fig. 5, lane 4) showed that the change had similar effect as the Phe20 to Stop mutant. This data supports our hypothesis that the encoded peptide is the key element.

Dr Pentecost showed earlier that a prox-uORF that opened normally but terminated at residue three was not inhibitory, indicating that the act of translational initiation was not the sole critical factor. This was extended by introducing stop codons at a variety of residues. Initially we scanned the sequences but now have a fairly comprehensive analysis for the C-terminal half of the uORF.

A composite summary of our results is included in Fig. 6. The data show some clear trends: The N-terminal regions of the uORF peptide are not the critical feature and can be eliminated. From introduced stops it appears that loss of the residues C-terminal to Phe13 results in loss of inhibitory action. Loss of the C-terminal Phe 20 and Gly19-Phe20 results in an increase inhibition. The C terminal half of the peptide is apparently the most important region but this has to be qualified by noting the earlier observation that distal uORF did not cause statistically significant inhibition of downstream ER-GFP expression while the two uORFs share a common final 5 codons. Some of the data indicates a role for Arg15, the final residue specific to proximal transcripts. It may be that association of this residue with the common C-terminal sequence is a factor that leads to statistically significant inhibition of expression by uORF region the proximal but not distal ER promoter transcripts.

N-terminally shortened uORFs are clearly active as inhibitors. These studies used the Phe 20 truncated sequence with the uORF eliminated (see [5] in Fig 2). The uORF encoded in this sequence would be potent but lacks an initiator. The reintroduction of a start in place of Leu 10 or Arg 15 restored inhibition (Figs. 4 &

6). This data indicates a critical feature in the residues 16-18 region. The Trp16-Pro17-Ala18 sequences is clearly the core of the translational inhibitory motif, though adjacent residues have influence. There is some risk in using the uORF lacking the C-terminal Phe20 as the basic standard for introducing new translational starts. However, the constructs lacking the C-terminal Phe20 have the big advantage of giving a strong effect so that changes in expression are easily examined.

We attempted further alanine scanning mutagenesis in the codon 15-20 region. Replacement of Phe 20 by alanine resulted decreased reporter expression (significant at the 1% level) and mutation of Pro17 caused even greater inhibition (Fig. 8) Mutation of Trp 16 marginally increased inhibition while Arg15 had no effect. This data again suggests C-terminal residues are important but the data is frankly contradictory. There are two possible explanations for this: firstly with some changes there could be noise from effects on expression beyond the translational control level and secondly we have gone below the resolution of our technique - multiple residues in the C-terminal region likely contribute to a structural motif. We intend to carry out simultaneous mutagenesis of several residues to alanine as an approach to identify key motifs. Our focus is on the interesting Trp-Pro combination and flanking residues though Pro17 will not be a target itself as the Pro 17 to Ala change showed strong inhibition of downstream expression).

Parallel analysis of RNA and protein expression: Our hypothesis is that differing expression levels from alternate ER-GFP constructs are due to altered protein yield caused by translational control. We therefore predict that equivalent amount of mRNA from constructs containing uORF (with/without different mutations) and the ER-ORF have different capacities to make protein.

The studies described above measure changes in expressed reporter but do not truly address the source of changes. In order to test the hypothesis, we have begun a series of experiments that assess relative levels of RNA expression in addition to protein expression in transiently transfected cells.

Our experimental approach is to measure the level of ER-GFP RNA and protein expression and to normalize this to the co-transfected HEM45 which is expected to be constant except for effects resulting from variation in transfection efficiency. Protein expression is measured by western blot as before while RealTime quantitative PCR assays are used to measure the level of expressed RNA for each GFP fusion. A cartoon of the PCR system is shown in Fig. 9. The reverse primer is to common GFP sequences. This primer only binds sequence from the transfected construct as GFP is not normally expressed in mammalian cells. Specificity for the reporter sequence and to the invariant reference is through the forward primers designed to bind ER and HEM45 sequences respectively. The test constructs include that with wild type uORF, a uORF ATG to CTG mutant (eliminates uORF translational initiation), C-terminal Phe20 to a termination codon mutant (exhibits enhanced inhibition), and a double mutant containing both the ATG to CTG mutation and C-terminal Phe20 to a termination codon mutation (Fig. 2, cartoons 1- 4).

RNA from cells transiently transfected with the four ER-GFP HEM45-GFP constructs was prepared using Triazol (Invitrogen, Life Sci). After quantification of total RNA by measuring abosrption at 260nm, we used real time RT-PCR kit to quantify the level for target mRNAs using standard curves made from existing plasmid constructs. These curves were stored and used throughout this study to allow calculation of relative changes in target mRNAs.

PCR assays utilized the Light Cycler (Roche, Indianapolis, IN) together with the Qiagen (Valencia, CA) One-Step Kit supplemented with Syber Green I (10,000× stock, Molecular Probes, Eugene, OR). Master mixes were prepared at 4°C and in multiples of 50 µl containing: 33 µl water; 10 µl of 5× buffer (supplied); 2 µl of dNTP solution (supplied); 2 µl of enzyme mixture (supplied); 1 µl of primer mixture (25 µM each); and 2 µl of Syber Green I (diluted 1/5000 with water). Aliquots (14.3 µl) were distributed into capillaries precooled to 4°C, and total RNA was added in 0.75 µl. The capillaries were spun just prior to RT-PCR. The synthesis of cDNA was performed at 50°C for 30 min followed by a 15 min at 95°C heating step to inactivate the reverse transcriptases and activate the Taq polymerase. The number of amplification cycles was 45, and the cycle profile was denaturation (95°C, 15 s), annealing (60°C, 15 s) and extension (72°C, 30 s). A standard

RNA was included with each test series samples to compensate for run to run variation and relate data to the stored standard curves.

Primer sets amplifying fragments specific for ER-GFP or HEM45-GFP were designed using Primer 3 PCR primer selection program (Whitehead Institute, MIT). The forward primer for amplifying a fragment in ER-GFP (ER/Forward) is 5'CATGACCCTCCACACCAAA3'. The forward primer for amplifying a fragment in HEM45-GFP (HEM45/Forward) is 5'CAAGAGCATCCAGAACAGCC3'. The reverse primer (GFP/Reverse) is 5'AAGTCGTGCTGCTTCATGTG 3'.

Quantitation of bands in western blots is difficult as film does not have a linear response. To avoid this problem we made a serial dilution of one sample (ATG-CTG mutation Fig. 10, Lanes 1-6) and estimated fold changes by matching test samples band intensity to bands in the dilution and then using the *dilution factor* for quantitation.

The data shows that when we loaded equal amounts of the test samples (tracks 6-8) there was equal expression of the HEM45-GFP reference but changing expression of the test ER-GFP cassette, as expected from prior data. The ER-GFP cassette is identical in all the test constructs but lies downstream of the putative regulatory uORF/uORF mutants.

RNA data is shown in Fig 11. Normalized RNA expression was equal, within a factor of two, for the proximal uORF construct (track 9 to match Fig. 10) and the mutant lacking the terminal phenylalanine residue (track 8) while protein expression changed at least four-fold (Fig. 10) supporting the relevance of translational control to this system. Data where we eliminated the uORF also supports translational control but is less convincing. The constructs lacking the translational start (ATG to CTG mutation) are showing enhanced RNA levels relative to the normalizing HEM45-GFP construct which complicates analysis. There was a >16-fold change in protein expression for the Phe20 to stop mutant uORF construct as compared to the equivalent construct lacking the uORF translational start. There was a change in RNA expression but this was only three-fold. There was an ~eight-fold change in protein expression with a three- fold change in RNA expression for the construct carrying a normal uORF as compared to expression from a construct where the uORF translational start was eliminated. We have a variety of other constructs were mutations abolish uORF inhibitory effects (e.g. termination of uORF at codon three) and these will be tested to see if they provide a simpler comparison for analysis.

Task 2. *Vary the spacing between the main ER translational start and the terminator of the prox uORF. Decrease and increase the spacing by looping in/out regions using modified PCR primers. Carry out studies using either GFP fusions or use the entire ER as reporter.*

This topic was adressed in the year 1 report. The following material is excerpted from the year 1 report.
A series of constructs varying spacing between the uORF and ER main translational starts have been made in the context of our ER-GFP construct. 15, 30 and 40 nt sequences were looped out in the 52 nt inter-cistronic space between the uORF and the main ER translational start, adjacent to the uORF terminator (see Fig.12 cartoon). The effect on ER-GFP expression was assessed by western blot analysis. Results were equivocal (see blot data in Fig 12) as GFP expression from the construct with the largest deletion was *increased* compared to the parental ER-GFP construct while the expression from the others was l*ower*.

We have not followed up on the data from year 1 as spacing between the uORF and main ER translational start does not appear crucial to uORF action in the ER prox-promoter transcript. Interpretation of these studies will always potentially difficult as mechanism could be based on the separation of the two translational starts or could be related to the need for translational re-initiation prior to translation of the down stream ER-GFP. These two events are related, but mechanisms could exist that are more related to one aspect rather than the other.

Task 3A. *Generate expression constructs for full length ER protein containing 5' regions of both the proximal and distal constructs, introduce to ER negative cells and assess relative efficiency of ER expression using western analysis, determine extent of regulatory effects on ER levels in comparison to, say, estradiol treated cells (estrogen tends to destabilize its receptor). Include, as controls, similar constructs where the uORF translational starts have been eliminated (ATG to AAG mutation). Extend studies to ER regulated reporter constructs i.e. use the modified ER expression constructs to modulate expression of luciferase constructs carrying estrogen response elements.*

We have created ER expression constructs in pRCCMV that express the human ER ORF of pHEO (11) but where the 5' regions are those of the transcript or are modified from the wild type sequence. The two modifications were to eliminate the uORF by mutation of the initiator codon to CTG and, separately, to eliminate the terminal Phe20codon. Mutagenesis was by the Quikchange system and utlized primers used for earlier studies. Essentially these are the constructs of cartoons in Fig. 2 but now the downstream expressed cassette is the intact ER ORF. ER expression was monitored by two assays: western blot (not shown) and action on a co-transfected reporter..

The reporter construct for measuring ER action was a firefly luciferase expression construct carrying an estrogen response element in the promoter region. Expression was normalized to expression of a cotransfected Renilla luciferase expression construct luciferases were sequentially assayed using the Promega dual luciferase assay. We have transfected cells with varying amounts of the constructs and treated cells with a standard amount (10nM) estradiol. A control group receiving no ER expression plasmid was normalized to 1.0. Studies are still in progress with this assay. Changes appear limited (Fig. 13), activity in cells transfected with the more inhibitory uORF construct (Phe20 to stop) was consistently lower than the original construct and the latter trended lower than the construct were the uORF translational start was eliminated. The model assay faces several challenges: Firstly there is a need to optimize amounts of transfected ER plasmid as we know that too much expression plasmid leads to reduced ER-related gene activation and the need to optimize estradiol level in that context. Treatment of cells with estradiol results in ER turnover/degradation so that the act of trying to assay the ER by reporter (by loading receptor with hormone) can lead to changes. It actually is possible that the translational control does not show up with estradiol treated receptor but might be found with alternate activation mechanism. The question underlying our studies is how does alternate promoter in tumors affect prognosis and we know that most of breast tumors are in older post-menopausal women who will not experience cyclic spiking in circulatory 17βestradiol levels. No difference was seen with the alternate expression constructs in western immunoblots. A fundamental concern is that we have to transfect microgram amounts of DNA vectors to detect ER in westerns while nanogram amounts are optimal for activating reporter expression.

Task 3B. *Finally, time permitting, determine if the regulatory effects of the proximal uORF are similar in both normal immortalized lines (eg MCF-10) and breast tumor lines (eg MDA-MB-231).*

Several constructs were studied in cell lines additional to Hela. The basic ER-GP construct and the matching construct where the prox-uORF was eliminated by mutation of the translational start were introduced to MCF-7 cells (ER positive breast cancer line principally utilizing ER prox promoter), MDA MB 231 (ER Negative breast cancer cells) and Ishikawa (ER positive uterine adenocarcinoma utilizing both the prox and distal promoters). This data was introduced in task 1 (Fig. 5). The ER GFP constructs were used so that we could assay effects in cell environments that have ER and because of the limited success at studying uORF action on full-length ER.

We found that elimination of the uORF increased expression from the downstream ER-GFP ORF; similar to behavior in Hela. Likewise an increased inhibition of downstream expression was found in all tested cell lines upon elimination of the final Phenylalanine residue of the uORF. This was relieved by elimination of the uORF translational start again supporting the hypothesis of translational control by the uORF peptide.

Key Accomplishments:

* Demonstrated/confirmed translational regulation of estrogen receptor expression from the proximal promoter transcript by the transcript specific upstream open reading frame

* Defined a core sequence responsible for affecting ER expression.

* Established that regulation was similar in multiples cell types

* Developed methods for normalizing of GFP protein in western blots of transfected cells by using cotransfection with a larger GFP derivative that is detected by GFP antisera

* Developed methods for quantitating GFP RNA expression and normalizing for transfection efficiency using cotransfection a larger GFP derivative.

Reportable Outcomes

1. Application for Predoctoral Fellowship (BC021821), CDMRP BCRP : Translational Control of Estrogen Receptor Alpha (ER): Mechanism for Action by a Peptide Encoded Upstream of ER in the ER Proximal Promoter Transcript, June, 2002. Not funded

2. Application for Predoctoral Fellowship (BC030820), CDMRP BCRP : Translational Control of Estrogen Receptor Alpha (ER): Mechanism for Action by a Peptide Encoded Upstream of ER in the ER Proximal Promoter Transcript, May, 2003

3. Translational Regulation of Estrogen Receptor Alpha Expression, P27-28, Fall 2002 Era of Hope Meeting, Luo& Pentecost, abstract appended.

Supported Personnel

Luo (family name), Mingfei (given name): Graduate research assistant

Conclusions

The studies have largely confirmed the starting hypothesis that there is translational regulation of estrogen receptor expression from the proximal promoter transcript by the transcript specific upstream open reading frame. We have defined a core sequence Trp-Pro-Ala in the C-terminal region of the uORF peptide that is responsible for this effect and shown that flanking residues affect the action of this core sequence.

Our studies now are largely focused on two related topics: 1. Structural features of the uORF peptide that makes it a translational inhibitor and 2. The mechanism by which the presence of the uORF results in translational inhibition.

We are currently seeking to disrupt the uORF action on downstream expression *without* affecting RNA expression and to then assess protein expression. We have several choices to investigate, such as any of the uORFs terminating prior to the C-terminal Arg 15-Trp16-Pro17-Ala18 region. We may then directly investigate the structure of the peptide and inactive mutants using biophysical approaches. Based upon our data and studies of other uORFs by other labs we are designing further studies based on the hypothesis that the action of the uORF is during the act of translation on the ribosome. If the hypothesis is true then one would expect effects in all cells plus we can test the hypothesis in cell free reticulocyte translation systems. We aim to address how the peptide works by studying ribosome association with transcripts. We can use 'toe-printing' methodologies to determine If the ribosomes stall on the transcript in the region of the uORF.

References

1. **Osborne, C.K.** 1998 Steroid hormone receptors in breast cancer management. Breast Cancer Research & Treatment 51:227-238
2. **Leclercq, G.** 2002 Molecular forms of the estrogen receptor in breast cancer. J Steroid Biochem Mol Biol 80:259-272
3. **Kozak, M.** 1987 At least six nucleotides preceding the AUG initiator codon enhance translation in mammalian cells. Journal of Molecular Biology 196:947-950
4. **Kozak, M.** 1989 The scanning model for translation: an update. Journal of Cell Biology 108:229-241
5. **Hinnebusch, A.G.** 1996 Translational Control of GCN4: Gene-specific regualtion by Phosphorylation of eIF2. Translational Control.CSH Laboratory Press,199-244
6. **Muckenthaler, M., Gray, N.K., Hentze, M.W.** 1998 IRP-1 binding to ferritin mRNA prevents the recruitment of the small ribosomal subunit by the cap-binding complex eIF4F. Mol Cell 2:383-388
7. **Geballe ,A.P.** 1996 Translational Control Mediated by Upstream AUG Codons. In: Hershey J.W.B, Matthews MB, Sonenberg N (eds). Translational control.CSH,173-197
8. **Fasco, M.J.** 1997 Quantitation of estrogen receptor mRNA and its alternatively spliced mRNAs in breast tumor cells and tissues. Analytical Biochemistry 245:167-178
9. **Pentecost, B.T.** 1998 Expression and estrogen regulation of the HEM45 MRNA in human tumor lines and in the rat uterus. J Steroid Biochem Mol Biol. 64(1-2):25-33
10. **Fasco, M. J., Amin, A., Pentecost, B.T., Yang, Y., Gierthy, J.F.**, Changes in Estrogen Receptor-Related Phenotypic Changes in MCF-7 Cells During Prolonged Exposure to Tamoxifen, Accepted Mol Cell Endocrinol
11. **Green, S., Walter, P., Kumar, V., Krust, A., Bornert, J.M., Argos, P. and Chambon, P.** Human oestrogen receptor cDNA: sequence, expression and homology to v-erb-A Nature 320: 134-139 (1986)

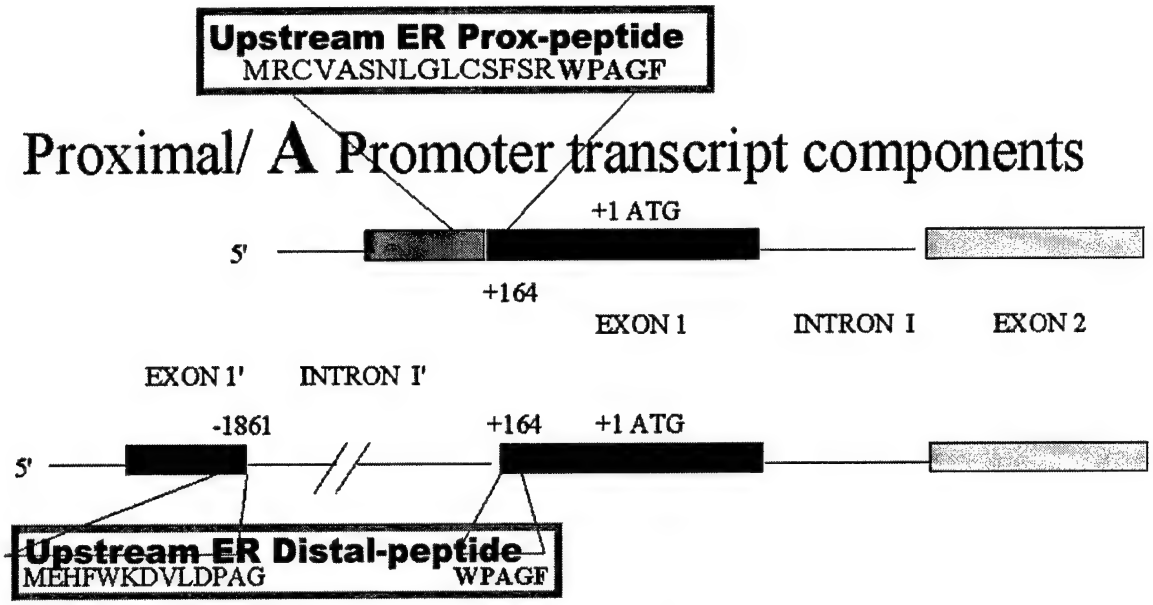


Fig. 1. Organization of the 5' regions of the human ER-alpha gene: Two promoters (distal [B, lower] and proximal [A, upper]) are present in the ER- gene and utilized in reproductive tissues. Alternative promoter transcripts converge at nt 164 of the mature mRNA (numbering based on proximal promoter transcript) and have a common ER coding region. Short ORF's are present in the upstream 5' region, these differ in the their N-terminal regions but share 5 C-terminal residues, by (RNA) splicing. Peptide sequences are shown

Region	DNA Sequence and Potential Translational Product
Distal 5' uORF	MetGluHisPheTrpLysAspValLeuAspProAlaGly**TrpProAlaGlyPheTer** ..caagcccATGGAACATTTCTGGAAAGACGTTCTTGATCCAGCAGGG**TGGCCCGCCGGTTTCTGA**gcc...
Proximal 5' uORF	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 MetArgCysValAlaSerAsnLeuGlyLeuCysSerPheSerArg**TrpProAlaGlyPheTer** ..gcgggacATGCGCTGCGTCGCCTCTAACCTCGGGCTGTGCTCTTTTTCCAG**GTGGCCCGCCGGTTTCTGA**gcc.
Major ER ORF	MetThrMetThrLeuHisThrLysAlaSerGlyMetAlaLeuLeuHisGlnIle .... ccacggaccATGACCATGACCCTCCACACCAAAGCATCTGGGATGGCCCTACTGCATCAGATCCA...
Kozak Optimal Translational Start	Met ... gccgccRccATGG...

Fig. 1B: Translational Start Regions of Estrogen Receptor Open Reading Frames

The DNA sequences and entire potential encoded peptides of alternate ER transcript upstream regions are shown together with the fragment of the main ER coding region that is included in ER-GFP fusion constructs. The DNA sequences 5' of the translational starts are shown and are included in constructs as they contribute to the relative strength of translational start sites. The five shared codons (from a common exon) of the proximal and distal transcripts are shown in bold. For comparison the optimal translational start defined by Kozak is included; the critical elements of the ATG flanking sequences are the purine (R) at -3 and the guanine residue at +4.

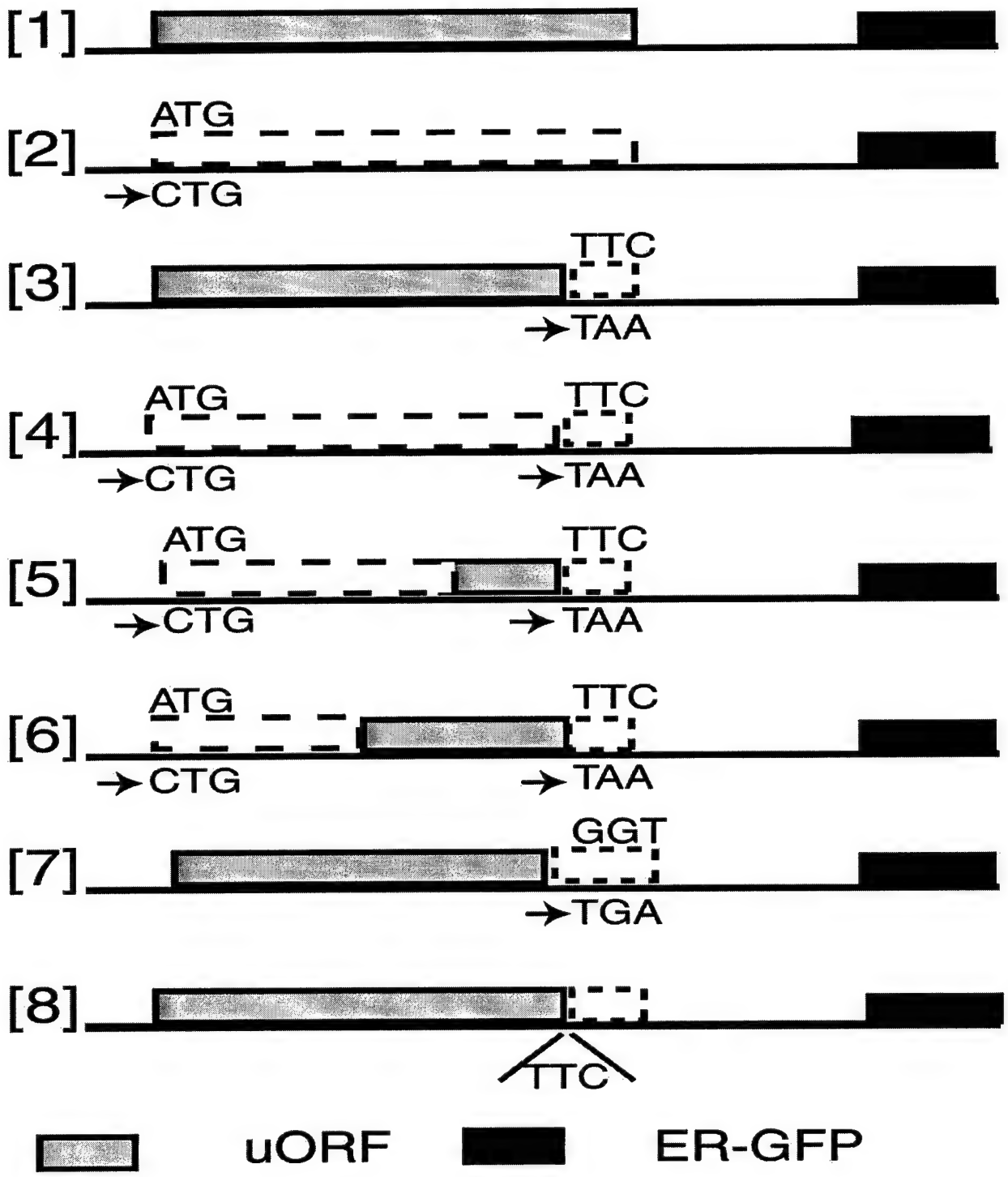


Fig.2. Cartoon of GFP constructs

Construct [1]: containing wild type proximal ER upstream regions and main ER ORF sequence (18 AA) -GFP fusion.
Construct [2]: same as [1] except the uORF translational start ATG is eliminated by mutation to CTG
Construct [3]: same as [1] except uORF C-terminal Phe mutated to a termination codon (TTC->TAA)
Construct [4]: same as [3] except uORF translational start ATG is eliminated by mutation to CTG
Construct [5]: same as [4] except introduced ATG at codon 15
Construct [6]: same as [4] except introducted ATG at codon 10
Construct [7]: same as [1] except uORF C-terminal Gly mutated to a termination codon (GGT->TGA)
Construct [8]: same as [1] except uORF C-terminal Phe is deleted (looped out)

Proximal 5' uORF

```
                 1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20
               MetArgCysValAlaSerAsnLeuGlyLeuCysSerPheSerArgTrpProAlaGlyPheTer
     ..gcgggacATGCGCTGCGTCGCCTCTAACCTCGGGCTGTGCTCTTTTTCCAGGTGGCCCGCCGGTTTCTGAgcc.
Arg->Gln  (BP491/2)
      CGCGGGACATGCAGTGCGTCGCCTCTAAC
Leu->Ter  (BP584/5)            CTAACCTCGGGTAGTGCTCTTTTTCCAGG
Cys->Stop (BP484/5)                CCTCGGGCTGTGATCTTTTTCCAGGT
Cys->Ser  (BP484/5)                CCTCGGGCTGTCCTCTTTTTCCAG
Ser->Ter  (BP586/7)                 CTCGGGCTGTGCTGATTTTCCAGG
Phe->Ter  (BP592/3)                    GGGCTGTGCTCTTGATCCAGGTGGCCC
Ser->Ter  (BP590/1)                       CTGTGCTCTTTTTGAAGGTGGCCCGCC
Arg->Ala  (BP493/4)                         GTGCTCTTTTTCCGCGTGGCCCGCCGG
Arg->Ter  (BP564/5)                         GTGCTCTTTTTCCTGATGGCCCGCCGGTTTC
Trp->Ala  (BP495/6)                              CTTTTTCCAGGGCGCCCGCCGGTTTCTG
Trp->Ter  (BP502/3)                            GCTCTTTTTCCAGGTGACCCGCCGGTTTC
Pro->Ala  (BP489/90)                             CTTTTTCCAGGTGGGCCGCCGGTTTC
Pro->Ter  (BP566/7)                              CTTTTTCCAGGTGGTGAGCCGGTTTCTGAG
Ala->Ter  (BP568/9)                                    CCAGGTGGCCCTGAGGTTTCTGAGCC
Gly->Ter  (BP518/9)                                        GTGGCCCGCCTGATTCTGAGCCTTCTGC
Phe->Ter  (BP504/5)                                        GTGGCCCGCCGGTTGATGAGCCTTCTGC
          (BP512/3)                                        GTGGCCCGCCGGTTAATGAGCCTTCTGC
Phe->Ala  (BP516/7)                                        GTGGCCCGCCGGTGCCTGAGCCTTCTGC
Phe->Tyr  (BP514/5)                                        GTGGCCCGCCGGTTACTGAGCCTTCTGC
```

Double Mutant

Arg->Gln (BP491/2) and Second Arg->Ala (BP493/4)

```
               MetArgCysValAlaSerAsnLeuGlyLeuCysSerPheSerArgTrpProAlaGlyPheTer
     ..gcgggacATGCGCTGCGTCGCCTCTAACCTCGGGCTGTGCTCTTTTTCCAGGTGGCCCGCCGGTTTCTGAgcc.
      CGCGGGACATGCAGTGCGTCGCCTCTAAC        GTGCTCTTTTTCCGCGTGGCCCGCCGG
```

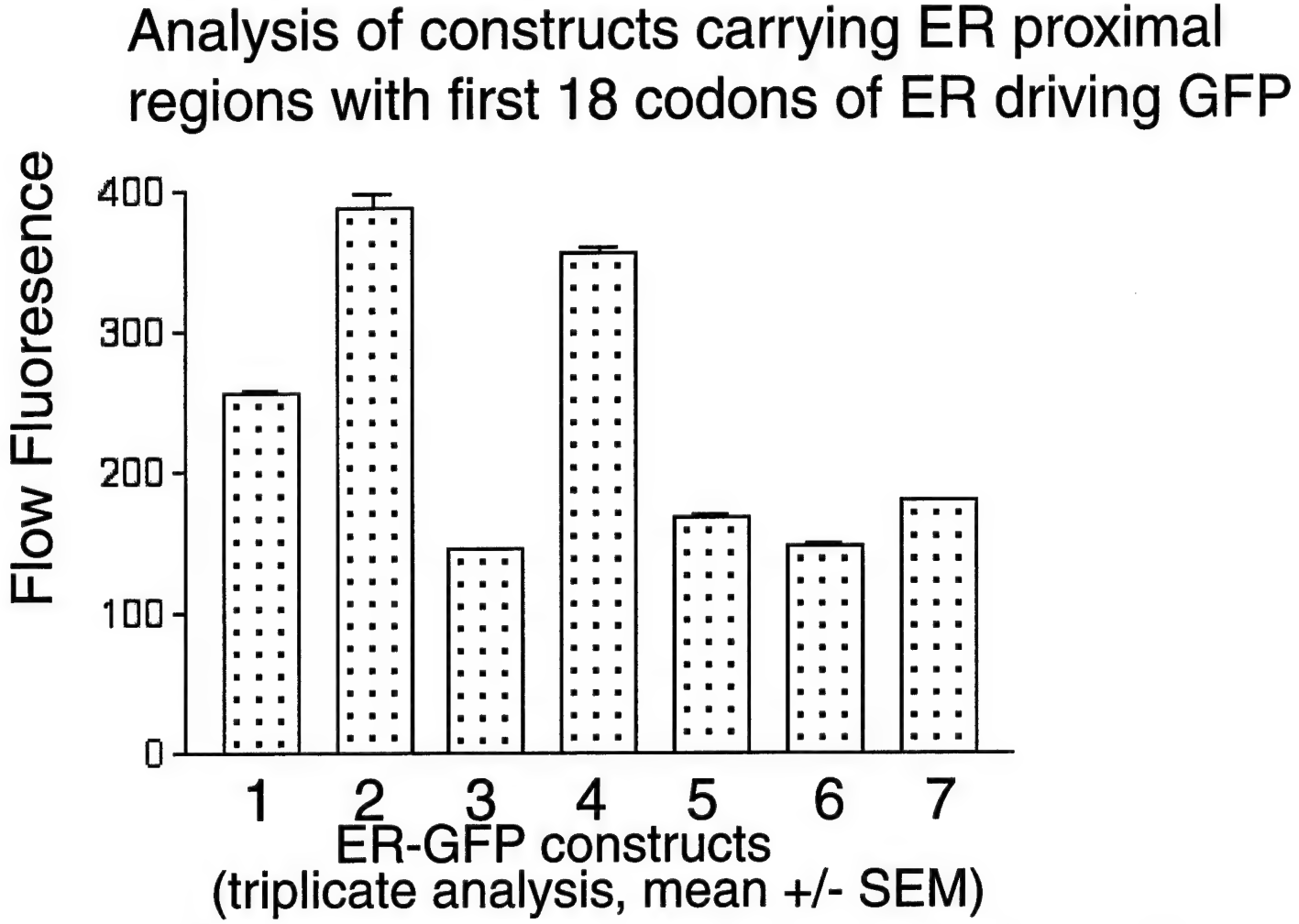
Spacing mutants, ER prox ORF separation from main ER start

```
Minimal separation, ~10bp left (main Kozak region) 5'CCGCCGGTTTCTGAGCCACGGACCATGAC3'
~15bp separation removed, by uORF stop             5'CCGCCGGTTTCTGAGGGACACGGTCTGCA3'
~30bp separation removed, by uORF stop             5'CCGCCGGTTTCTGACCCTGCCCGCGGCCAC3'
```

All the primers shown above are forward primer from 5' to 3'. The reverse primers are complementary to the listed forward primers.

Fig. 3: Summary of oligonucleotides utilized in this project

uORF peptide target/ consequences are described in left column.
DNA target is illustrated in the main panel.



1.Proximal ER-GFP construct (ER proximal transcript, 7 bp 5' of the uORF ATG to 56 bp 3' of ATG including 1st 18 codons , fused in frame to GFP) This is construct [1] in Fig 2.

2. Proximal ER-GFP construct as 1, but uORF ATG-->CTG (construct [2] in Fig 2)

3. Proximal ER-GFP construct as 1, but uORF last codon was mutated from Phe (TTC) to Stop (TAA) (this is construct [3] in Fig 2.)

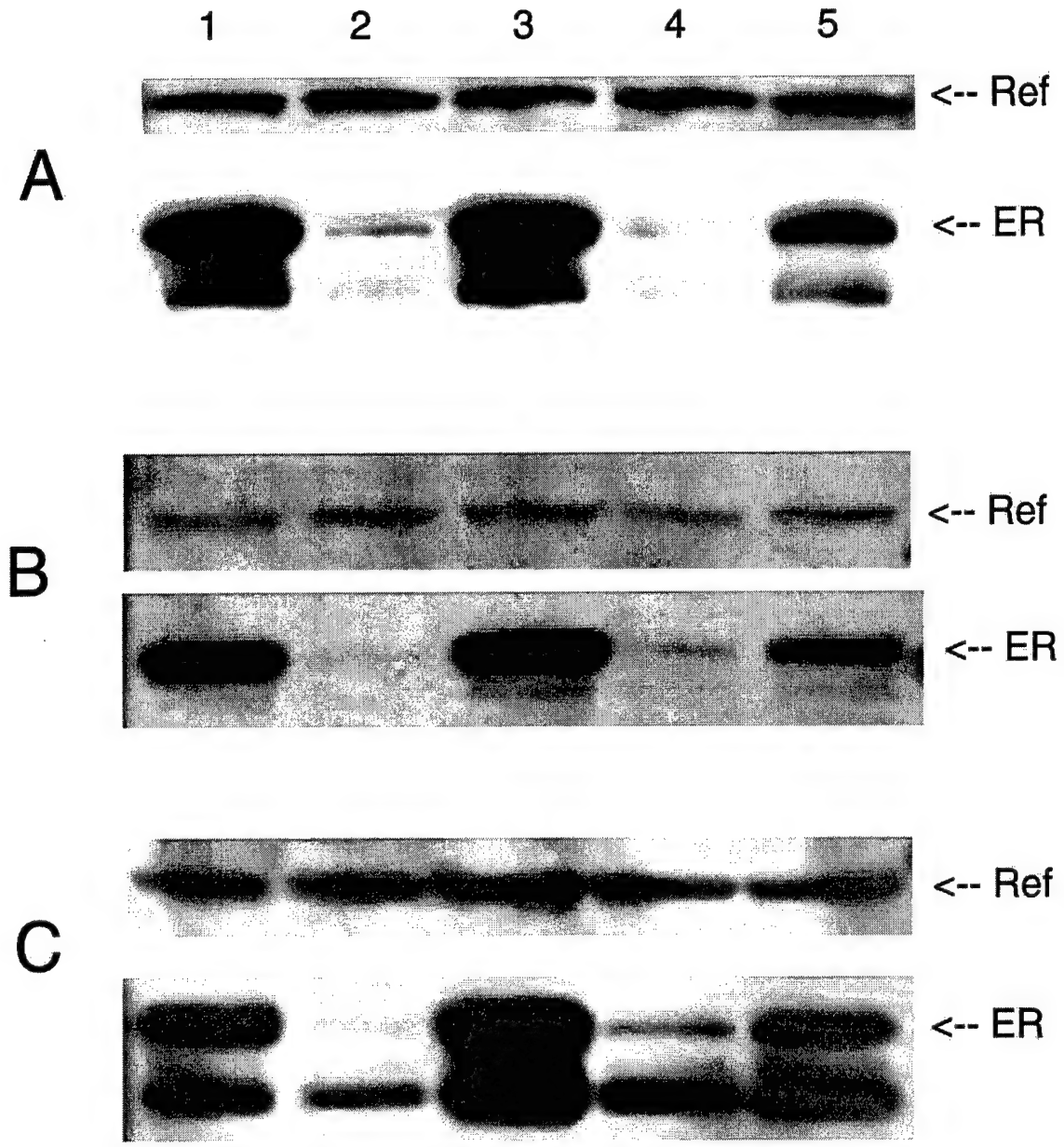
4. Proximal ER-GFP construct as 4, but uORF ATG-->CTG, and last Phe ->STOP ([4] in Fig .2)

5. Proximal ER-GFP construct as 4, but uORF ATG-->CTG, and ATG introduced at codon 15 ([5] in Fig.2)

6. Proximal ER-GFP construct as 4, but uORF ATG-->CTG. and ATG introducted at codon 10 ([6] in Fig.2)

7. Proximal ER-GFP construct as 4, but uORF codon Gly19 was mutated to Stop (TGA) ([7] in Fig.2)

Fig. 4. ER proximal uORF transcript sequences modify expression of an ER-GFP reporter

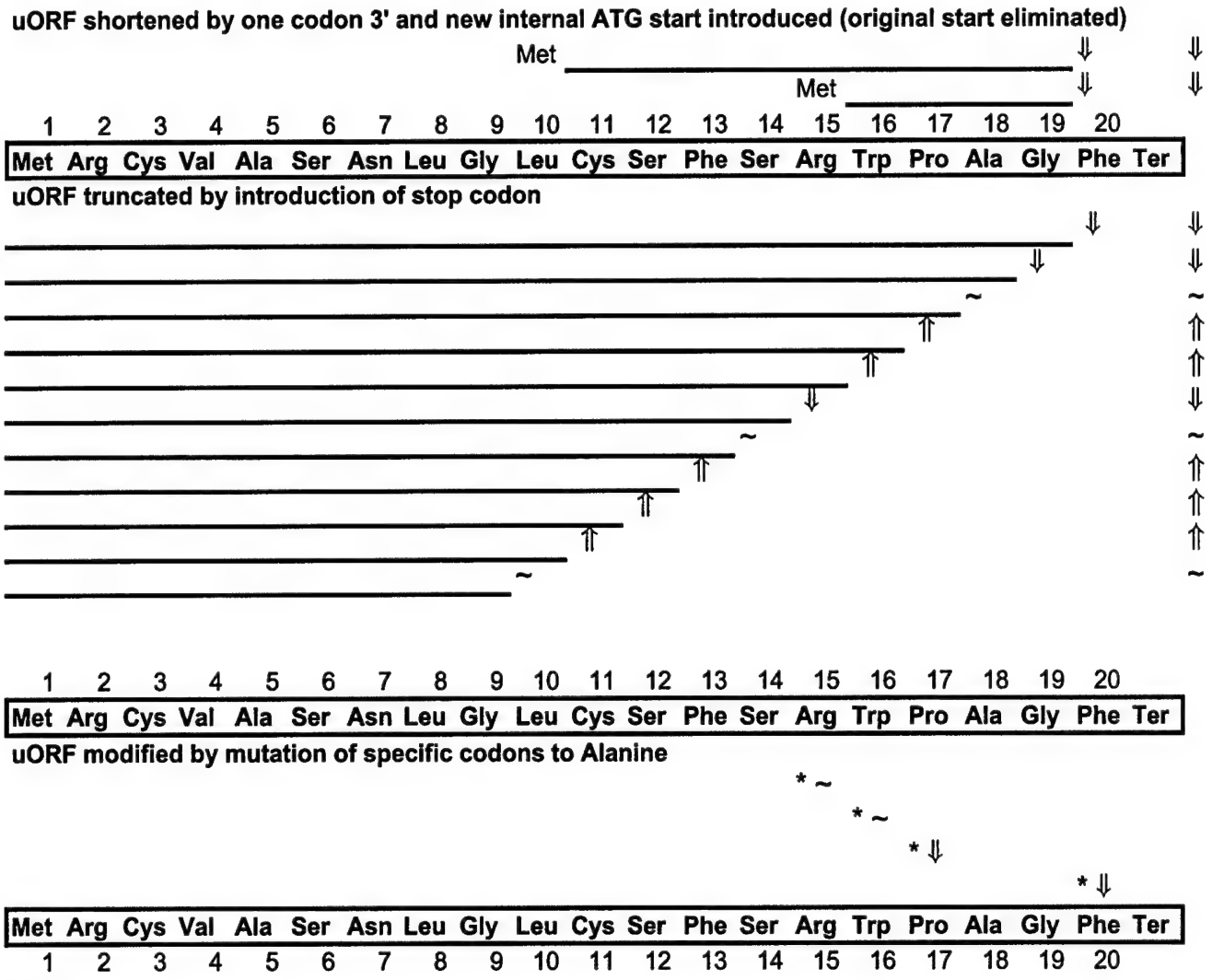


A: MDA-MB 231 B: MCF-7 C: ISHIKAWA

1. Proximal ER-GFP mt (uORF ATG-->CTG, **[2]** in Fig.2)
2. Proximal ER-GFP mt (uORF last Phe-->Stop, **[3]** in Fig.2)
3. Proximal ER-GFP mt (uORF ATG-->CTG and last Phe->Stop, **[4]** in Fig.2)
4. Proximal ER-GFP mt (uORF last Phe is looped out, **[8]** in Fig.2)
5. Proximal ER-GFP (**[1]** in Fig.2)

Fig. 5. Western blot analysis of ER-GFP expression in transiently transfected Ishikawa, MCF-7 & MDA-MB231

Reference signal is generated from cotransfected invariant HEM45-GFP construct.



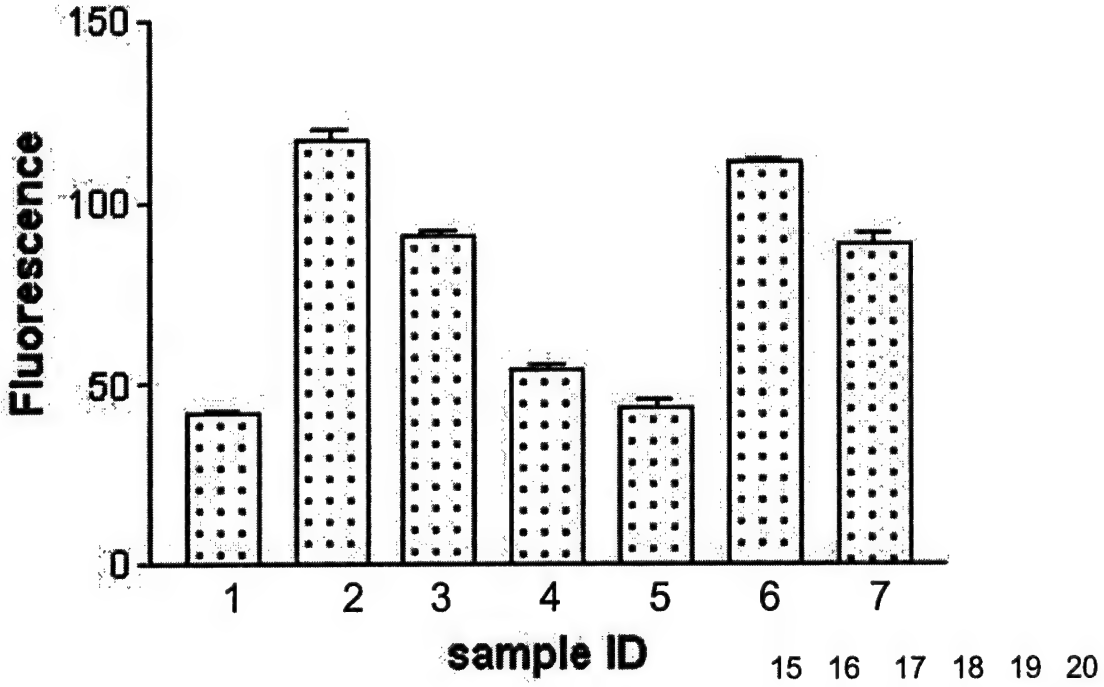
Key:

~ Downstream protein expression level is aproximately the same as the wild type

⇓ Downstream protein expression level is decreased

⇑ Downstream protein expression level is increased

Fig. 6: Summation of effects of uORF sequence modifications on downstream ER-GFP reporter expression

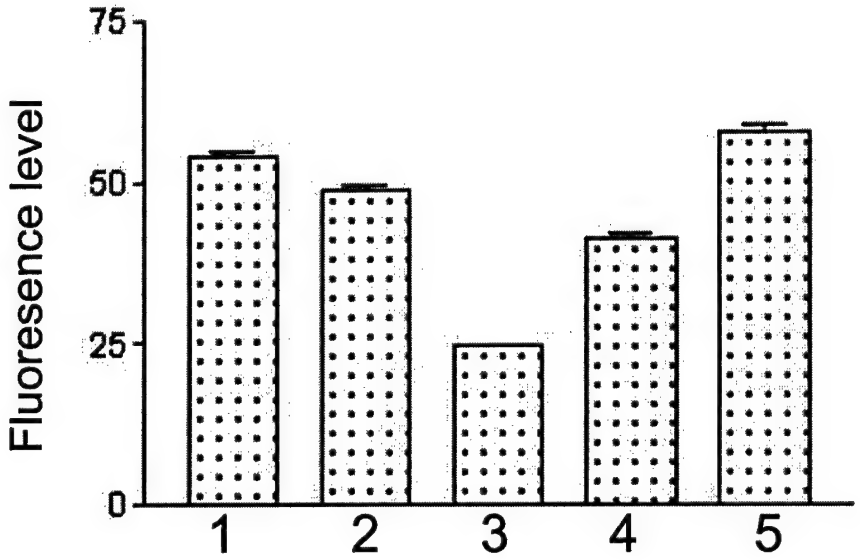


MetArgCysValAlaSerAsnLeuGlyLeuCysSerPheSerArgTrpProAlaGlyPheTer

1. Arg15-Stop
2. Pro17-Stop
3. Ala18-Stop
4. Gly19-Stop
5. Phe 20-Stop
6. Double mutant (Phe-Stop and uORF ATG->CTG, [4] in Fig. 2)
7. Basic unmodified Prox ER-GFP construct (see [1] in Fig. 2)

Fig. 7: Introduction of termination codons in the prox uORF DNA sequence affects downstream ER-GFP expression variably

Effects of DNA sequence mutations were determined by flow-cytometry on transiently transfected HeLa cells.



MetArgCysValAlaSerAsnLeuGlyLeuCysSerPheSer
ArgTrpProAlaGly**Phe**Ter
15 16 17 18 19 20

1. p3005-5 (Arg15-Ala) P value 0.1-0.15
2. p3006-A (Trp16-Ala) P value 0.01-0.05
3. p3003-3 (Pro17-Ala) P value <0.0005
4. p3016-A (Phe20-Ala) P value 0.005-0.01
5. p2483-3 Basic unmodified Prox ER-GFP construct (see [1] in Fig. 2)

Fig. 8: Mutation to Alanine of specific codons in the prox uORF DNA sequence affects downstream ER-GFP expression variably

Effects of DNA sequence mutations were determined by flow-cytometry on transiently transfected HeLa cells.

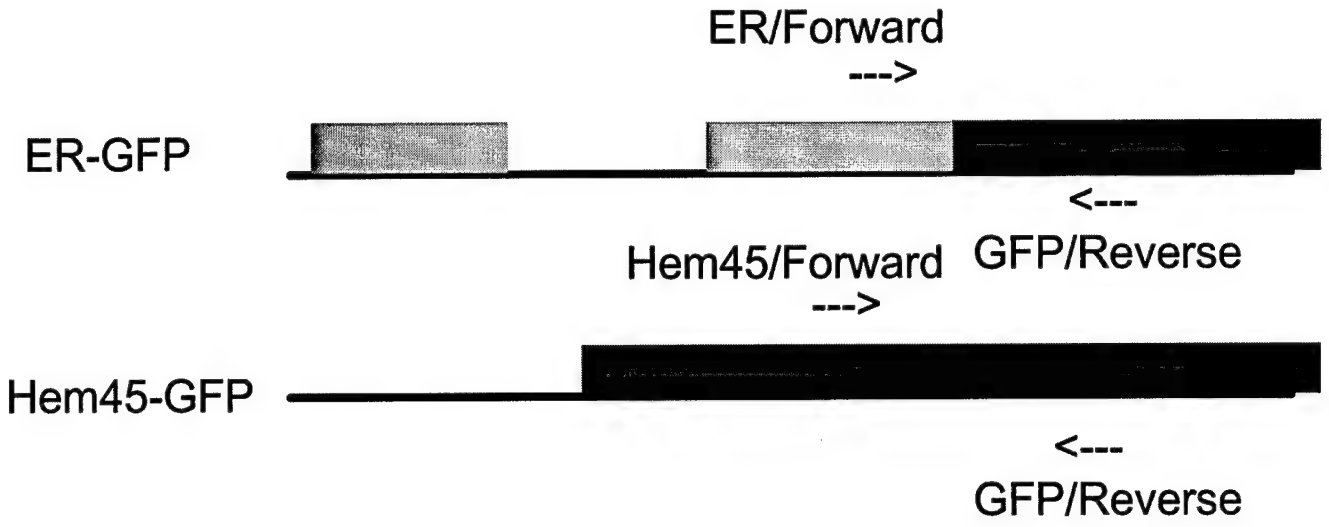
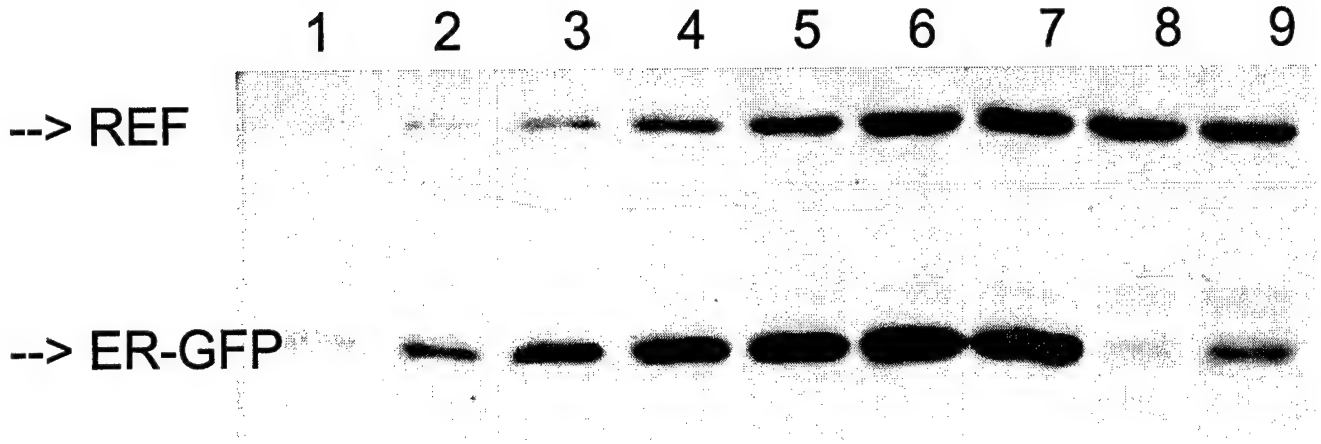


Fig. 9: Cartoon of ER-GFP and Hem45-GFP constructs and relative position of primers amplifying a fragment specific for ER-GFP or Hem45-GFP.

Levels of GFP RNA expression in transfected cells wa determined quantitative PCR in the Roche LightCycler. ER-GFP levels were determined in separate reactions using a common reverse primer an target specific forward primers



Lane 1-6 are serial dilution of sample from the proximal ER-GFP mt (uORF ATG->CTG, **[2]** in Fig. 2)
1. 0.625ug Hela cell protein, 2. 1.25ug, 3. 2.5ug,
4. 5.0ug, 5. 10ug
6. 20ug ER-GFP mt (uORF ATG->CTG, **[2]** in Fig. 2)
7. Proximal ER-GFP mt (uORF ATG-> CTG and Phe 20-> Stop, **[4]** in Fig. 2)
8. Proximal ER-GFP mt (uORF Phe20 -> Stop, **[3]** in Fig. 2)
9. Basic unmodified Proximal ER-GFP construct (**[1]** in Fig. 2)

Fig. 10: Western blot analysis of ER-GFP expression in transiently transfected Hela cells using anti-GFP antisera

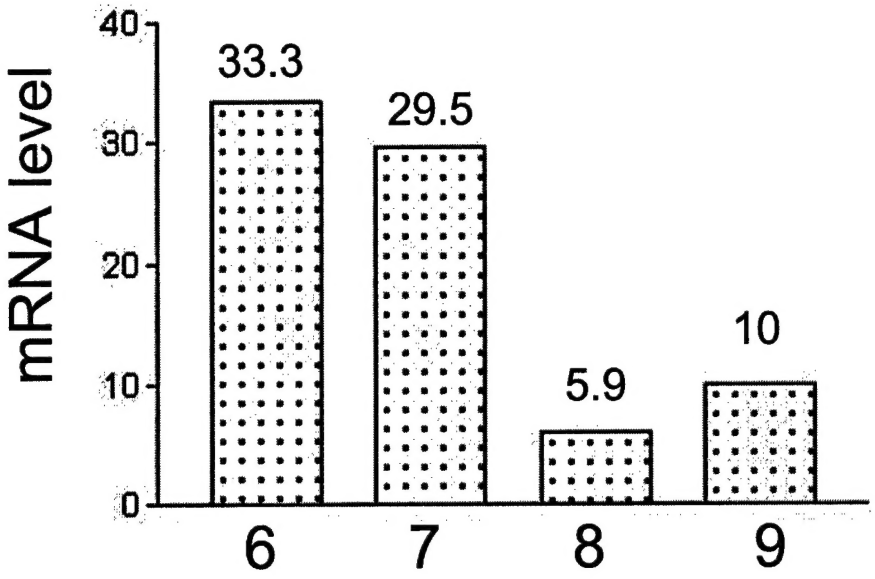


Fig. 11 RT-PCR analysis of ER-GFP mRNA levels in transiently transfected Hela cells.

Samples in this Fig. are numbered to match test tracks in the Fig. 10 Western blot

6. Proximal ER-GFP mt (uORF ATG->CTG, **[2]** in Fig. 2)
7. Proximal ER-GFP mt (uORF ATG-> CTG and uORF Phe 20-> Stop, **[4]** in Fig. 2)
8. Proximal ER-GFP mt (uORF Phe 20-> Stop, **[3]** in Fig. 2)
9. Basic unmodified Proximal ER-GFP construct (**[1]** in Fig. 2)

The number on the top bars represents the relative mRNA level given normalized to a vaue of 10 for the basic unmodified Proximal ER-GFP construct. Relative mRNA levels are the levels ER-GFP mRNA normalized to HEM45-GFP mRNA levels.

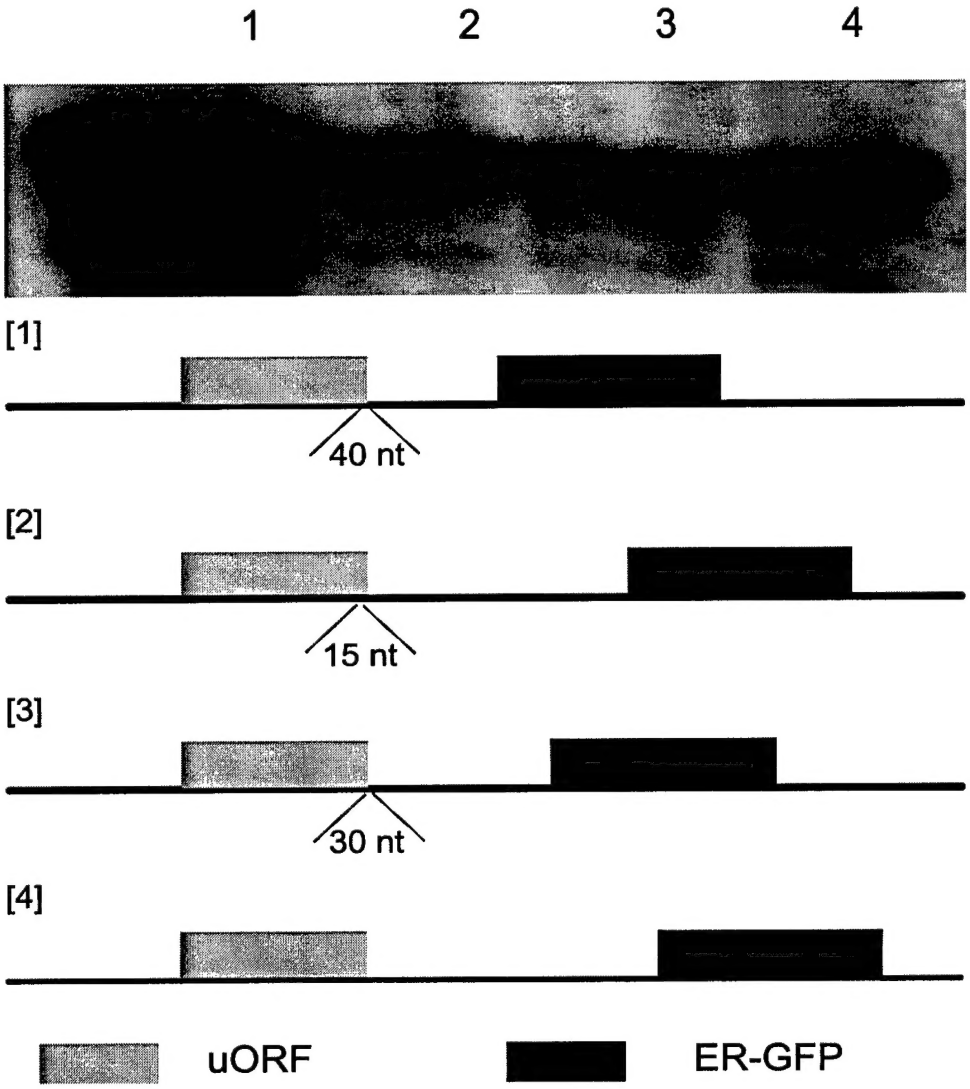


Fig. 12: Variable effects of changing the separation of the prox uORF and main ER translational start

Top panel shows western blot analysis of ER-GFP expression in transiently transfected cells. Rabbit polyclonal GFP antibody is used and the protein is resolved by 12% polyacrylamide gel. Lower panel is a cartoon of GFP constructs. Construct [1]: same as [4] except the space between proximal ER upstream region and main ER ORF is decreased by 40 nts Construct [2]: same as [4] except the space between proximal ER upstream region and main ER ORF is decreased by 15 nts Construct [3]: same as [4] except the space between proximal ER upstream region and main ER ORF is decreased by 30 nts Construct [4]: contain wild type proximal ER upstream regions and main ER ORF sequence (18 AA)-GFP fusion. The uORF terminates 52 nt 5' to the main ER start

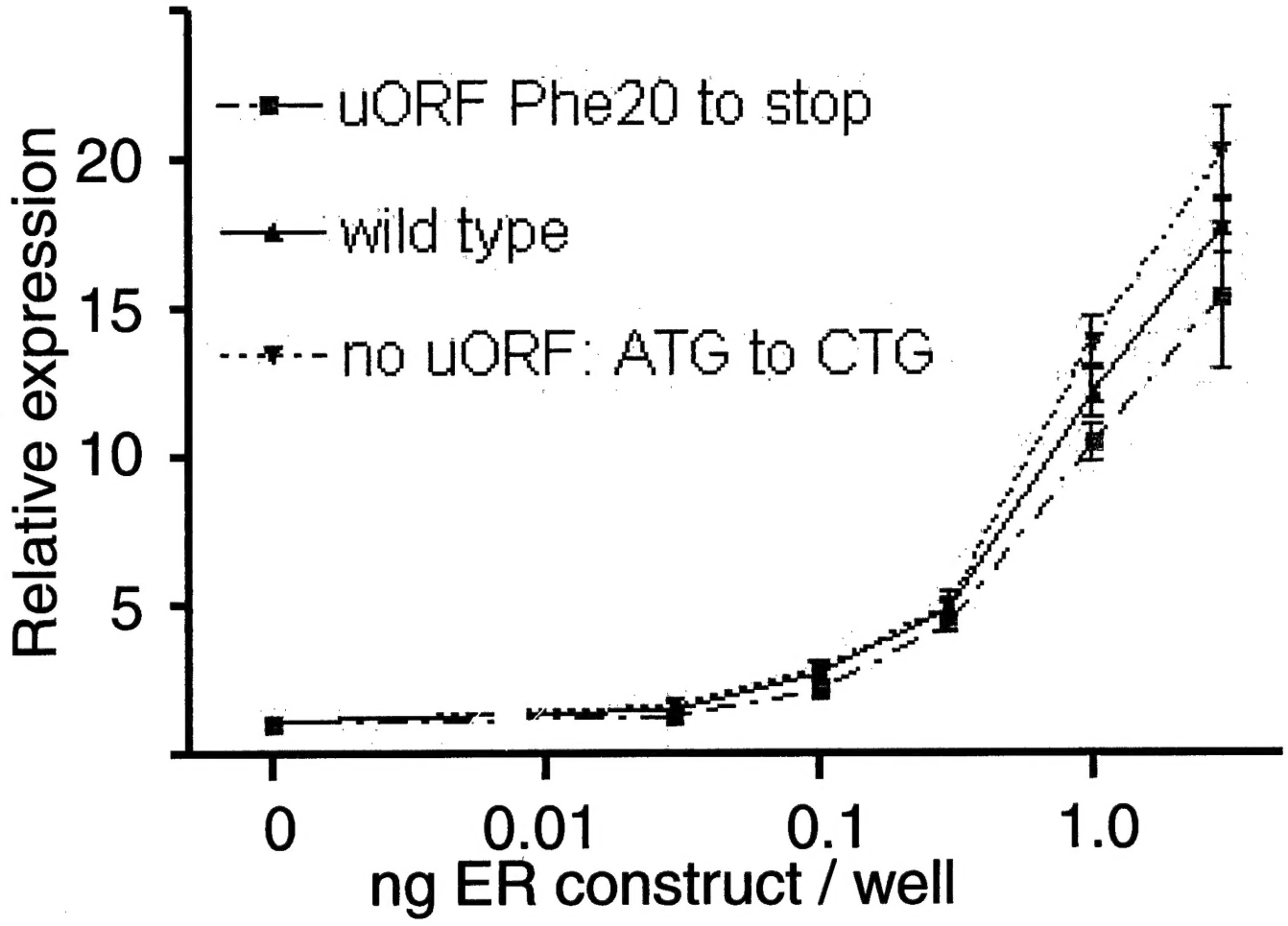


Fig. 13: Limited effects of uORF modifications on an Estrogen responsive reporter

A luciferase reporter was used to test the effect of uORFs on ER-based gene activation. Cells were transfected with lipofectamine 2000 essentially as described (10). Increasing amounts of ER expression constructs were included. All cells were treated with estradiol at 10nM and cells harvested after 48h. There is a trend for the cultures receiving construct lacking a uORF to have increased reporter expression and for the construct with a shortened uORF to show reduced reporter expression.
We are trying to optimize conditions in order to achieve statistically significant changes.

TRANSLATIONAL REGULATION OF ESTROGEN RECEPTOR ALPHA EXPRESSION

Luo, Mingfei, and Pentecost, Brian T.

Wadsworth Center, New York State Department of Health,
Albany, NY

pentecos@wadsworth.org

Estrogen Receptor (ER)-mRNA is transcribed from two promoters, variable relative use occurs in breast tumors. The proximal (prox) ER promoter transcribes a complete exon-one whereas the distal promoter is 2kb upstream in the genome and splices into nt164 of the prox form. The 5' region of ER prox promoter transcript contains a 20 amino acid (aa) residue upstream open reading frame (uORF) besides the main ER ORF. The distal promoter transcript contains an 18 aa residue uORF, sharing five in-frame codons with the prox-uORF; uORFs terminate 50 bases upstream of the main ER translational start which is shared by both transcripts.

Using flow cytometry and western blot techniques we have shown that eliminating the prox-uORF increases the expression from ER-Green Fluorescent Protein (GFP) fusion protein constructs of prox-ER transcript upstream sequences and the first 18 ER codons. We hypothesize the reported lack of correlation between ER-protein and prox-transcript levels in breast tumors (Hayashi *et al* 1997 Carcinogenesis **18** 459-464) is due to inhibitory translational control by the prox-peptide. Replacing the C-terminal (Phe) codon with termination-codons (UAA or UGA) further inhibited expression of ER-GFP fusion proteins, however constructs in which the new C-terminal codon was removed had a significantly greater level of GFP expression. The super inhibitory effect of the Phe to stop mutations seems related to the peptide of the ORF as suppression effects were reversed by eliminating the prox-uORF translational start codon. This data suggests a key role for C-terminal residues. We introduced ATG translational starts within the abolished prox-uORF; the truncated ORFs were less effective in inhibiting translation, suggesting a role for multiple residues beyond the C-terminal region. Effects were similar in each of several ER positive and negative cell lines.

The mechanism of translational inhibition by the prox-uORF likely involves interactions of encoded peptide sequences, rather than being simply due to upstream initiation and the subsequent need for re-initiation at the main ER translational start. Differences in properties of the non-shared regions of the distal transcript peptide could account for the lack of statistically significant inhibition by the uORF in distal ER transcript-GFP constructs. Our model is that the prox-uORF peptides inhibit the progression of ribosomes at the time of synthesis; the effect is seen in all cell lines and action of the ORF peptides should be specific to prox-ER transcripts.

The goals of our studies are to understand the control of ER expression, to determine consequences of alternate promoter and possibly develop strategies to manipulate ER expression based on alternate promoter features.

The U.S. Army Medical Research and Materiel Command under DAMD17-01-1-0529 supported this work.